

WHAT IS CLAIMED IS

1. An isolated or recombinant polypeptide with an improved endo-protease activity selected from among increased thermotolerance, increased activity at ambient temperature, increased
5 activity at alkaline pH, increased activity at acid pH, increased activity at neutral pH, and increased activity in the presence of organic solvents, relative to the subtilisin homologue polypeptide corresponding to SEQ ID NO: 261, which polypeptide has at least 70% sequence identity to at least one of SEQ ID NO: 131 to SEQ ID NO: 260, over a comparison window of at least 20 contiguous amino acids.
10
2. The isolated or recombinant polypeptide of claim 1, which polypeptide has at least 80% sequence identity to at least one of SEQ ID NO: 131 to SEQ ID NO: 260 over a comparison window of at least 20 contiguous amino acids.
- 15 3. The isolated or recombinant polypeptide of claim 1, which polypeptide has at least 85% sequence identity to at least one of SEQ ID NO: 131 to SEQ ID NO: 260 over a comparison window of at least 20 contiguous amino acids.
- 20 4. The isolated or recombinant polypeptide of claim 1, which polypeptide has at least 90% sequence identity to at least one of SEQ ID NO: 131 to SEQ ID NO: 260 over a comparison window of at least 20 contiguous amino acids.
- 25 5. The isolated or recombinant polypeptide of claim 1, which polypeptide has at least 95% sequence identity to at least one of SEQ ID NO: 131 to SEQ ID NO: 260 over a comparison window of at least 20 contiguous amino acids.
6. The isolated or recombinant polypeptide of claim 1, which polypeptide has at least 96% sequence identity to at least one of SEQ ID NO: 131 to SEQ ID NO: 260 over a comparison window of at least 20 contiguous amino acids.
30
7. The isolated or recombinant polypeptide of claim 1, which polypeptide has at least 97% sequence identity to at least one of SEQ ID NO: 131 to SEQ ID NO: 260 over a comparison window of at least 20 contiguous amino acids.

8. The isolated or recombinant polypeptide of claim 1, which polypeptide has at least 98% sequence identity to at least one of SEQ ID NO: 131 to SEQ ID NO: 260 over a comparison window of at least 20 contiguous amino acids.
- 5 9. The isolated or recombinant polypeptide of claim 1, which polypeptide has at least 99% sequence identity to at least one of SEQ ID NO: 131 to SEQ ID NO: 260 over a comparison window of at least 20 contiguous amino acids.
- 10 10. The isolated or recombinant polypeptide of claim 1, which polypeptide comprises an amino acid sequence of any one of SEQ ID NO: 131 to SEQ ID NO: 260 over a comparison window of at least 20 contiguous amino acids.
- 15 11. The isolated or recombinant polypeptide of claim 1, which polypeptide comprises an amino acid sequence of any one of SEQ ID NO: 131 to SEQ ID NO: 260 over a comparison window of at least 30 contiguous amino acids.
- 20 12. The isolated or recombinant polypeptide of claim 1, which polypeptide comprises an amino acid sequence of any one of SEQ ID NO: 131 to SEQ ID NO: 260 over a comparison window of at least 50 contiguous amino acids.
13. The isolated or recombinant polypeptide of claim 1, which polypeptide comprises an amino acid sequence of any one of SEQ ID NO: 131 to SEQ ID NO: 260 over a comparison window of at least 100 contiguous amino acids.
- 25 14. The isolated or recombinant polypeptide of claim 1, which polypeptide comprises an amino acid sequence of any one of SEQ ID NO: 131 to SEQ ID NO: 260.
- 30 15. The isolated or recombinant polypeptide of claim 1, which polypeptide comprises at least 30 contiguous amino acid residues of a polypeptide encoded by a coding polynucleotide sequence, the coding polynucleotide sequence selected from the group consisting of:
- (a) a polynucleotide selected from any of SEQ ID NO: 1 to SEQ ID NO: 130;
 - (b) a polynucleotide sequence that encodes a polypeptide selected from any of SEQ ID NO: 131 to SEQ ID NO: 260; and

(c) a polynucleotide sequence which hybridizes under stringent conditions over substantially the entire length of a polynucleotide sequence (a) or (b).

16. The isolated or recombinant polypeptide of claim 1, wherein identity is determined by a sequence alignment performed using BLASTP with default parameters set to measure 70% identity.

17. An isolated or recombinant polypeptide with endo-protease activity, which endo-protease activity comprises a conditional property selected from among pH dependence, temperature dependence, dependence on ionic strength, activation by ligand binding, and inactivation by ligand binding, relative to the subtilisin homologue polypeptide corresponding to SEQ ID NO: 261, which polypeptide has at least 70% sequence identity to at least one of SEQ ID NO: 131 to SEQ ID NO: 260 over a comparison window of at least 20 contiguous amino acids.

18. The isolated or recombinant polypeptide of claim 17, which polypeptide has at least 80% identity to at least one of SEQ ID NO: 131 to SEQ ID NO: 260 over a comparison window of at least 20 contiguous amino acids.

19. The isolated or recombinant polypeptide of claim 17, which polypeptide has at least 90% identity to at least one of SEQ ID NO: 131 to SEQ ID NO: 260 over a comparison window of at least 20 contiguous amino acids.

20. The isolated or recombinant polypeptide of claim 17, which polypeptide has at least 95% identity to at least one of SEQ ID NO: 131 to SEQ ID NO: 260 over a comparison window of at least 20 contiguous amino acids.

21. The isolated or recombinant polypeptide of claim 17, which polypeptide has at least 99% identity to at least one of SEQ ID NO: 131 to SEQ ID NO: 260 over a comparison window of at least 20 contiguous amino acids.

22. The isolated or recombinant polypeptide of claim 17, which polypeptide comprises an amino acid sequence of any one of SEQ ID NO: 131 to SEQ ID NO: 260 over a comparison window of at least 20 contiguous amino acids.

23. The isolated or recombinant polypeptide of claim 17, which polypeptide comprises an amino acid sequence of any one of SEQ ID NO: 131 to SEQ ID NO: 260 over a comparison window of at least 30 contiguous amino acids.

5 24. The isolated or recombinant polypeptide of claim 17, which polypeptide comprises an amino acid sequence of any one of SEQ ID NO: 131 to SEQ ID NO: 260 over a comparison window of at least 50 contiguous amino acids.

10 25. The isolated or recombinant polypeptide of claim 17, which polypeptide comprises an amino acid sequence of any one of SEQ ID NO: 131 to SEQ ID NO: 260 over a comparison window of at least 100 contiguous amino acids.

26. The isolated or recombinant polypeptide of claim 17, which polypeptide comprises an amino acid sequence of any one of SEQ ID NO: 131 to SEQ ID NO: 260.

15 27. The isolated or recombinant polypeptide of claim 17, which polypeptide comprises at least 30 contiguous amino acid residues of a polypeptide encoded by a coding polynucleotide sequence, the coding polynucleotide sequence selected from the group consisting of:

- 20 (a) a polynucleotide selected from any of SEQ ID NO: 1 to SEQ ID NO: 130;
- (b) a polynucleotide sequence that encodes a polypeptide selected from any of SEQ ID NO: 131 to SEQ ID NO: 260; and
- (c) a polynucleotide sequence which hybridizes under stringent conditions over substantially the entire length of a polynucleotide sequence (a) or (b).

25 28. An isolated or recombinant polypeptide encoded by a nucleic acid of comprising a polynucleotide sequence selected from the group consisting of:

- (a) a polynucleotide sequence selected from SEQ ID NO: 1 to SEQ ID NO: 130, or a complementary polynucleotide sequence thereof;
- 30 (b) a polynucleotide sequence encoding a polypeptide selected from SEQ ID NO: 131 to SEQ ID NO: 260, or a complementary polynucleotide sequence thereof;
- (c) a polynucleotide sequence which hybridizes under highly stringent conditions over substantially the entire length of polynucleotide sequence (a) or (b), or which hybridizes to a subsequence thereof encoding at least 25 amino acids, which subsequence corresponds to

positions 71 through 95, 86 through 110, 111 through 135 or 196 through 220 of the mature Savinase® polypeptide;

(d) a polynucleotide sequence comprising all or a fragment of (a), (b), or (c), wherein the fragment encodes a polypeptide with endo-protease activity;

(e) a polynucleotide sequence encoding a polypeptide, the polypeptide comprising an amino acid sequence which is substantially identical over the entire length of any one of SEQ ID NO: 131 to SEQ ID NO: 260;

(f) a polynucleotide sequence encoding a polypeptide with an improved endo-protease activity selected from among increased thermotolerance, increased activity at ambient temperature, increased activity at alkaline pH, increased activity at acid pH, increased activity at neutral pH, and increased activity in the presence of organic solvents, relative to the subtilisin homologue polypeptide corresponding to SEQ ID NO: 261, which polypeptide has at least 70% sequence identity to a polynucleotide of (a), (b), (c) or (d); and

(g) a polynucleotide sequence encoding a polypeptide with an improved endo-protease activity selected from among increased thermotolerance, increased activity at ambient temperature, increased activity at alkaline pH, increased activity at acid pH, increased activity at neutral pH, and increased activity in the presence of organic solvents, relative to the subtilisin homologue polypeptide corresponding to SEQ ID NO: 261, produced by mutating or recombining one or more polynucleotide sequences of (a), (b), (c) or (d).

29. The isolated or recombinant polypeptide of claim 28, the polypeptide comprising an amino acid sequence of any one of SEQ ID NO: 131 to 260.

30. The polypeptide of claim 28, comprising an endo-protease.

31. The polypeptide of claim 30, comprising an endo-protease with at least one property selected from among increased thermotolerance, increased activity at ambient temperature, increased activity at alkaline pH, increased activity at acid pH, increased activity at neutral pH, and increased activity in the presence of organic solvents, relative to the subtilisin homologue polypeptide corresponding to SEQ ID NO: 261.

32. The polypeptide of claim 31, wherein the endo-protease comprises two or more properties selected from among increased thermotolerance, increased activity at ambient temperature, increased activity at alkaline pH, increased activity at acid pH, increased activity at

neutral pH, and increased activity in the presence of organic solvents, relative to the subtilisin homologue polypeptide corresponding to SEQ ID NO: 261.

33. The polypeptide of claim 28, the polypeptide comprising at least about 150 contiguous amino acids.

34. The polypeptide of claim 33, comprising a polypeptide with an endo-protease activity.

35. The polypeptide of claim 33, comprising at least 250 contiguous amino acids of the encoded protein.

36. The polypeptide of claim 35, comprising about 269 amino acids of the encoded protein, which encoded protein is a mature protein with endo-protease activity.

37. The polypeptide of claim 33, comprising at least 350 contiguous amino acids of the encoded protein.

38. The polypeptide of claim 37, comprising about 380 contiguous amino acids of the encoded protein, which encoded protein is a pre-pro peptide.

39. A polypeptide comprising the polypeptide sequence of any one of SEQ ID NO: 131 to SEQ ID NO: 260.

40. The polypeptide of claim 1, 17, 28 or 39, comprising a secretion/localization sequence.

41. The polypeptide of claim 1, 17, 28 or 39, further comprising a polypeptide purification subsequence.

42. The polypeptide of claim 41, wherein the polypeptide purification sequence is selected from the group consisting of: an epitope tag, a FLAG tag, a polyhistidine tag, and a GST fusion.

43. The polypeptide of claim 1, 17, 28 or 39, further comprising a Met at the N-terminus.

44. The polypeptide of claim 1, 17, 28 or 39, comprising a modified amino acid.

45. The polypeptide of claim 44, wherein the modified amino acid is selected from the group consisting of: a glycosylated amino acid, a PEGylated amino acid, a farnesylated amino acid, an acetylated amino acid, and a biotinylated amino acid.

5

46. A composition comprising the polypeptide of claim 1, 17, 28 or 39 and a detergent.

47. A polypeptide which comprises a unique subsequence in a polypeptide selected from: SEQ ID NO: 131 to SEQ ID NO: 260, wherein the unique subsequence is unique as compared to a polypeptide corresponding to any of: P29600, P41362, P29599, P27693, P20724, P41363, P00780, P00781, P35835, P00783, P29142, P04189, P07518, P00782, P04072, P16396, P29140, P29139, P08594, P16588, P11018, P54423, P40903, P23314, P23653, P33295, P42780, and P80146.

48. A polypeptide which is specifically bound by a polyclonal antisera raised against one or more antigen, the antigen comprising the sequence of SEQ ID NO: 131 to SEQ ID NO: 260, or a fragment thereof, wherein the antisera is subtracted with a naturally occurring subtilisin polypeptide or a subtilisin polypeptide corresponding to one or more of: P29600, P41362, P29599, P27693, P20724, P41363, P00780, P00781, P35835, P00783, P29142, P04189, P07518, P00782, P04072, P16396, P29140, P29139, P08594, P16588, P11018, P54423, P40903, P23314, P23653, P33295, P42780, and P80146.

49. An antibody or antisera produced by administering one or more antigen, which antigen comprises any one of SEQ ID NO: 131 to SEQ ID NO: 260, or a subfragment thereof, which antibody or antisera does not specifically bind to a naturally occurring subtilisin polypeptide or a subtilisin polypeptide corresponding to one or more of: P29600, P41362, P29599, P27693, P20724, P41363, P00780, P00781, P35835, P00783, P29142, P04189, P07518, P00782, P04072, P16396, P29140, P29139, P08594, P16588, P11018, P54423, P40903, P23314, P23653, P33295, P42780, and P80146.

30

50. An antibody or antisera which specifically binds a polypeptide, the polypeptide comprising a sequence selected from the group consisting of SEQ ID NO: 131 to SEQ ID NO: 260, wherein the antibody does not specifically bind to a naturally occurring subtilisin polypeptide or a subtilisin polypeptide corresponding to one or more of: P29600, P41362,

P29599, P27693, P20724, P41363, P00780, P00781, P35835, P00783, P29142, P04189, P07518, P00782, P04072, P16396, P29140, P29139, P08594, P16588, P11018, P54423, P40903, P23314, P23653, P33295, P42780, and P80146.

5 51. An isolated or recombinant nucleic acid, comprising a polynucleotide sequence selected from the group consisting of:

(a) a polynucleotide sequence comprising one or more of SEQ ID NO: 1 to SEQ ID NO: 130, or a complementary polynucleotide sequence thereof;

10 (b) a polynucleotide sequence encoding a polypeptide comprising a polypeptide sequence selected from the group of polypeptide sequences consisting of SEQ ID NO: 131 to SEQ ID NO: 260, or a complementary polynucleotide sequence thereof;

15 (c) a polynucleotide sequence which hybridizes under highly stringent conditions over substantially the entire length of polynucleotide sequence (a) or (b), or which hybridizes to a subsequence thereof comprising at least 75 nucleotides, which subsequence corresponds to positions encoding amino acids 71 through 95, 86 through 110, 111 through 135 or 196 through 220 of the mature Savinase® polypeptide; or

(d) a polynucleotide sequence comprising a fragment of (a), (b) or (c), which fragment encodes a polypeptide having endo-protease activity

20 52. The nucleic acid of claim 51, comprising a polynucleotide sequence encoding a polypeptide, the polypeptide comprising the amino acid sequence STQDGNGHGTHVAGT-X₇₀-AAL-X₇₄-N-X₇₆-X₇₇-GV-X₈₀-GVAP-X₈₅-X₈₆-X₈₇-LY-X₉₀-VKVL-X₉₅-A-X₉₇-G-X₉₉-GS-X₁₀₂-S-X₁₀₄-IA-X₁₀₇-GL-X₁₁₀-W-X₁₁₂-X₁₁₃-X₁₁₄-N-X₁₁₆-M-X₁₁₈-IAN-X₁₂₂-SLG-X₁₂₆-X₁₂₇-X₁₂₈-PS-X₁₃₁-TL-X₁₃₄-X₁₃₅-AVN-X₁₃₉-ATS-X₁₄₃-X₁₄₄-VLVIAA-X₁₅₁-GN-X₁₅₄-G-X₁₅₆-GSVGYPARYANA-MAVGATDQNN-X₁₇₉-RA-X₁₈₂-FSQYG-X₁₈₈-G-X₁₉₀-DIVAPGV-X₁₉₈-X₁₉₉-X₂₀₀-STYPG-X₂₀₆-X₂₀₇-Y-X₂₀₉-X₂₁₀-X₂₁₁-X₂₁₂-GTSMA-X₂₁₈-PHVAG-X₂₂₄-AAL, or a substituted variation thereof, wherein X₇₀ is I or V; X₇₄ is D or N; X₇₆ is D, S or N; X₇₇ is I, V or E; X₈₀ is I, V or L; X₈₅ is N, E or S; X₈₆ is A or V; X₈₇ is D or E; X₉₀ is A or G; X₉₅ is G, S or R; X₉₇ is S or N; X₉₉ is S, A or R; X₁₀₂ is I or V; X₁₀₄ is G or S; X₁₀₇ is R or Q; X₁₁₀ is E or Q; X₁₁₂ is A or S; X₁₁₃ is G or A, X₁₁₄ is E, A, T or N; X₁₁₆ is G or N; X₁₁₈ is D or H; X₁₂₂ is L or M; X₁₂₆ is S or T; X₁₂₇ is S or D; X₁₂₈ is A or F; X₁₃₁ is A, T or S; X₁₃₄ is E, K or G; X₁₃₅ is Q or R; X₁₃₉ is A or Y; X₁₄₃ is R or Q; X₁₄₄ is D or G; X₁₅₁ is S or T; X₁₅₄ is S or N; X₁₅₆ is A or S; X₁₇₉ is N or R; X₁₈₂ is S or N; X₁₈₈ is A or T; X₁₉₀ is L or I; X₁₉₈ is G, R or N; X₁₉₉ is V or L; X₂₀₀ is Q or R; X₂₀₆ is G, N, S or T; X₂₀₇ is R, S, T or Q; X₂₀₉ is V, A or D; X₂₁₀ is E, R or S; X₂₁₁ is L or M; X₂₁₂ is N, S or R; X₂₁₈ is S or T; and X₂₂₄ is A or V.

53. The nucleic acid of claim 51, which nucleic acid comprises a polynucleotide that encodes an endo-protease.
- 5 54. The nucleic acid of claim 51, which nucleic acid comprises a polynucleotide that encodes an endo-protease, which endo-protease is active at ambient temperature.
55. The nucleic acid of claim 51, which nucleic acid comprises a polynucleotide that encodes a thermotolerant endo-protease.
- 10 56. The nucleic acid of claim 55, which nucleic acid comprises a polynucleotide sequence selected from SEQ ID NO: 3, 7, 8, 10, 12, 14, 15, 16, 18, 21 and 25.
57. The nucleic acid of claim 51, which nucleic acid comprises a polynucleotide that encodes an endo-protease active at alkaline pH.
- 15 58. The nucleic acid of claim 57, which nucleic acid comprises a polynucleotide sequence selected from SEQ ID NO: 1, 17, 19, 22, 23, 24, 25, 26, 27 and 32.
- 20 59. The nucleic acid of claim 51, which nucleic acid comprises a polynucleotide that encodes an endo-protease active in an organic solvent.
60. The nucleic acid of claim 59, wherein the organic solvent comprises dimethylformamide (DMF).
- 25 61. The nucleic acid of claim 59, which nucleic acid comprises a polynucleotide sequence selected from SEQ ID NO: 2, 4, 5, 6, 11, 13, 20, 29, 30, and 33.
62. The nucleic acid of claim 51, which nucleic acid comprises a polynucleotide that encodes an endo-protease with two or more properties selected from among increased thermotolerance, increased activity at ambient temperature, increased activity at alkaline pH, increased activity at acid pH, increased activity at neutral pH, and increased activity in the presence of organic solvents, relative to the subtilisin homologue polypeptide corresponding to SEQ ID NO: 261.
- 30

63. An isolated or recombinant nucleic acid comprising a polynucleotide sequence encoding a polypeptide, the polypeptide comprising an amino acid sequence comprising at least 20 contiguous amino acids of any one of SEQ ID NO: 131-260.

5

64. The nucleic acid of claim 63, wherein the encoded polypeptide comprises at least about 50 contiguous amino acids of any one of SEQ ID NO: 131-260.

65. The nucleic acid of claim 63, wherein the encoded polypeptide comprises at least about 100 contiguous amino acids of any one of SEQ ID NO: 131-260.

10

66. The nucleic acid of claim 63, wherein the encoded polypeptide comprises at least about 150 contiguous amino acids of any one of SEQ ID NO: 131-260.

67. The nucleic acid of claim 63, wherein the encoded polypeptide is about 269 amino acids in length.

15

68. The nucleic acid of claim 63, wherein the encoded polypeptide is about 380 amino acids in length.

20

69. The nucleic acid of claim 68, wherein the encoded polypeptide is a pre-pro peptide.

70. An isolated or recombinant nucleic acid encoding a subtilisin, which subtilisin comprises a substitution of any one of SEQ ID NO: 131 to SEQ ID NO: 260 from amino acid 55 to amino acid 227 of the mature subtilisin protein.

25

71. The nucleic acid of claim 70, wherein the mature subtilisin protein is a mature Savinase® protein.

72. A composition comprising two or more nucleic acids of claim 51, 63 or 70.

30

73. A composition of claim 72, wherein the composition comprises a library comprising at least ten nucleic acids.

74. A composition produced by digesting one or more nucleic acid of claim 51, 63 or 70 with a restriction endonuclease, an RNase or a DNase.
- 5 75. A composition produced by incubating one or more nucleic acid of claim 51, 63 or 70 in the presence of deoxyribonucleotide triphosphates and a nucleic acid polymerase.
76. The composition of claim 75, wherein the nucleic acid polymerase is a thermostable polymerase.
- 10 77. A cell comprising at least one nucleic acid of claim 51, 63 or 70.
78. The cell of claim 77, wherein the cell expresses a polypeptide encoded by the nucleic acid.
- 15 79. A vector comprising the nucleic acid of claim 51, 63 or 70.
80. The vector of claim 79, wherein the vector comprises a plasmid, a cosmid, a phage or a virus.
- 20 81. The vector of claim 79, wherein the vector is an expression vector.
82. A cell transduced by the vector of claim 79.
83. A method of producing a nucleic acid encoding a subtilisin homologue polypeptide, the
25 method comprising recombining or mutating one or more nucleic acid of claim 51, 63 or 70 with one or more additional nucleic acid, the additional nucleic acid encoding a subtilisin homologue or subsequence thereof.
84. The method of claim 83, wherein the recombining comprises recursively recombining the
30 one or more nucleic acid.
85. The method of claim 83, wherein the recursively recombining is performed in vitro.
86. The method of claim 83, wherein the recursively recombining is performed in vivo.

87. The method of claim 83, wherein the recursively recombining produces at least one library of recombinant subtilisin homologue nucleic acids.

5 88. A nucleic acid library produced by the method of claim 86.

89. A population of cells comprising the library of claim 88.

90. A recombinant subtilisin homologue nucleic acid produced by the method of claim 83.

10

91. A cell comprising the nucleic acid of claim 90.

92. A method of producing a modified subtilisin nucleic acid homologue comprising mutating a nucleic acid of claim 51 or 63.

15

93. The modified subtilisin nucleic acid homologue produced by the method of claim 92.

94. A nucleic acid which comprises a unique subsequence in a nucleic acid selected from SEQ ID NO: 1 to SEQ ID NO: 130, wherein the unique subsequence is unique as compared to a nucleic acid corresponding to any of: M65086, D13157, S48754, AB005792, D29688, and M28537.

20

95. A target nucleic acid which hybridizes under stringent conditions to a unique coding oligonucleotide which encodes a unique subsequence in a polypeptide selected from SEQ ID NO: 131 to SEQ ID NO: 260, wherein the unique subsequence is unique as compared to a polypeptide corresponding to any of: : M65086, D13157, S48754, AB005792, D29688, and M28537.

25

96. A method of producing a polypeptide, the method comprising:

30

(a) introducing into a population of cells a nucleic acid of claim 51, 63 or 70, the nucleic acid operably linked to a regulatory sequence which directs expression of the encoded polypeptide

(b) culturing the cells in a culture medium to produce the polypeptide; and

(c) isolating the polypeptide from the cells or from the culture medium.

97. The method of claim 96, wherein the culturing is performed in a bulk fermentation vessel.

- 5 98. A method of producing a polypeptide, the method comprising:
- (a) introducing into a population of cells a recombinant expression vector comprising the nucleic acid of claim 51, 63 or 70;
 - (b) culturing the cells in a culture medium to produce the polypeptide encoded by the expression vector; and
 - 10 (c) isolating the polypeptide from the cells or from the culture medium.

99. The method of claim 98, wherein the culturing is performed in a bulk fermentation vessel.

- 15 100. A computer or computer readable medium comprising a database comprising a sequence record comprising one or more character string corresponding to a nucleic acid or protein sequence selected from SEQ ID NO: 1 to SEQ ID NO: 260.

20 101. An integrated system comprising a computer or computer readable medium comprising a database comprising one or more sequence records, each comprising one or more character string corresponding to a nucleic acid or protein sequence selected from SEQ ID NO: 1 to SEQ ID NO: 260, the integrated system further comprising a user input interface allowing a user to selectively view one or more sequence record.

25 102. The integrated system of claim 101, the computer or computer readable medium comprising an alignment instruction set which aligns the character strings with one or more additional character string corresponding to a nucleic acid or protein sequences.

30 103. The integrated system of claim 102, further comprising a user readable output element which displays an alignment produced by the alignment instruction set.

104. The integrated system of claim 101, the computer or computer readable medium further comprising an instruction set which translates one or more nucleic acid sequence comprising a sequence selected from SEQ ID NO: 1 to SEQ ID NO: 130, into an amino acid sequence.

105. The integrated system of claim 101, the computer or computer readable medium further comprising an instruction set for reverse-translating one or more amino acid sequence comprising a sequence selected from SEQ ID NO: 131 to SEQ ID NO: 260, into a nucleic acid sequence.

106. The integrated system of claim 105, wherein the instruction set selects the nucleic acid sequence by applying a codon usage instruction set or an instruction set which determines sequence identity to a test nucleic acid sequence.

107. A method of using a computer system to present information pertaining to at least one of a plurality of sequence records stored in a database, each comprising one or more character string corresponding to a nucleic acid or protein sequence selected from SEQ ID NO: 1 to SEQ ID NO: 260, the method comprising:

- (a) determining a list of one or more character strings corresponding to one or more of SEQ ID NO: 1 to SEQ ID NO: 260 or a subsequence thereof;
- (b) determining which character strings of the list are selected by a user; and
- (c) displaying the selected character strings, or aligning the selected character strings with an additional character string.

108. The method of claim 107, further comprising displaying an alignment of the selected character string with the additional character string.

109. The method of claim 108, further comprising displaying the list.

110. An isolated or recombinant polypeptide having endo-protease activity, which has an amino acid sequence that is at least 70% identical to at least one of SEQ ID NO: 131 to SEQ ID NO: 136 or SEQ ID NO: 138 to SEQ ID NO: 260.

111. The polypeptide of claim 110, which has an amino acid sequence that is at least 80% identical to at least one of SEQ ID NO: 131 to SEQ ID NO: 136 or SEQ ID NO: 138 to SEQ ID NO: 260.

112. The polypeptide of claim 110, which has an amino acid sequence that is at least 85% identical to at least one of SEQ ID NO: 131 to SEQ ID NO: 136 or SEQ ID NO: 138 to SEQ ID NO: 260.

5 113. The polypeptide of claim 110, which has an amino acid sequence that is at least 90% identical to at least one of SEQ ID NO: 131 to SEQ ID NO: 136 or SEQ ID NO: 138 to SEQ ID NO: 260.

10 114. The polypeptide of claim 110, which has an amino acid sequence that is at least 95% identical to at least one of SEQ ID NO: 131 to SEQ ID NO: 136 or SEQ ID NO: 138 to SEQ ID NO: 260.

15 115. The polypeptide of claim 110, which has an amino acid sequence that is at least 96% identical to at least one of SEQ ID NO: 131 to SEQ ID NO: 136 or SEQ ID NO: 138 to SEQ ID NO: 260.

20 116. The polypeptide of claim 110, which has an amino acid sequence that is at least 97% identical to at least one of SEQ ID NO: 131 to SEQ ID NO: 136 or SEQ ID NO: 138 to SEQ ID NO: 260.

117. The polypeptide of claim 110, which has an amino acid sequence that is at least 98% identical to at least one of SEQ ID NO: 131 to SEQ ID NO: 136 or SEQ ID NO: 138 to SEQ ID NO: 260.

25 118. The polypeptide of claim 110, which has an amino acid sequence that is at least 99% identical to at least one of SEQ ID NO: 131 to SEQ ID NO: 136 or SEQ ID NO: 138 to SEQ ID NO: 260.

30 119. The polypeptide of claim 110, which has an amino acid sequence that comprises one of SEQ ID NO: 131 to SEQ ID NO: 136 or SEQ ID NO: 138 to SEQ ID NO: 260.

120. A detergent composition comprising a polypeptide of claim 110 and a surfactant.

121. An isolated DNA sequence encoding a polypeptide of claim 110.

122. An expression vector comprising a DNA sequence of claim 121.

123. A microbial host cell transformed with an expression vector of claim 122.

5